NEW EXPERIMENTAL VALIDATION AND MODEL IMPROVEMENT TOOLS FOR THE CLIM2000 ENERGY SIMULATION SOFTWARE PROGRAM

Nadia RAHNI(*), Nacim RAMDANI(*), Yves CANDAU(*) and Gilles GUYON(**)
(*) LETIEF - Université PARIS XII - Val de Marne
94000 Créteil - France
(**)Electricité De France, Pôle Industrie, Division R&D, Site des Renardières
77818 Moret-sur-Loing Cédex, France - E-Mail : G.Guyon@edf.fr

ABSTRACT

This work has allowed to test different model improvement tools, by applying them on two building models.
At the close of this study, an important point concerning the capability of the CLIM2000 software program to perform exact derivative calculations came up : this advantage of the software make the sensitivity-uncertainty-optimisation work very accessible without increase of computer time.

INTRODUCTION

Within the context of a joint research project between Electricité de France and LETIEF, new experimental validation and numerical model improvement tools were developed with the CLIM2000 energy simulation software program in view. These tools are part of a global methodology aiming at understanding the changes in model behaviour when uncertain or inaccurate data (such as model parameters) vary.

The CLIM2000 energy simulation software program was designed at the Research Centre of Electricité De France - Les Renardières. This tool is operational since June 1989, and allows the behaviour of an entire building to be simulated. Its main objective is to produce economical studies, pertaining to energy balances over long periods as well as more detailed physical behaviour studies including stiff non-linear problems and varied dynamics.

CLIM2000 focuses on a modular vision of buildings, which, together with its equipment, are considered as an assembly of independent elements (doors, walls, windows, convectors, etc.) brought into relation by different thermal phenomena or control channels. Each element has a list of parameters which can be modified by the user, and is represented by an independent deterministic numerical model including the physical laws which characterise its behaviour. After the user has assembled these elementary models by means of a graphic interface, this produces a series of algebraic and differential equations which express the overall behaviour of the building.

These equations are then solved using the ESACAP solver, which is based on robust and efficient algorithms and which is able to efficiently process a great variety of applications: detailed analysis of comfort and heat exchanges, development of control and building energy management systems, energy and financial evaluation of different projects etc.

The experimental validation methodology is composed of three main stages, each of them being based on appropriate tools :

1- The model analysis, consisting in : a/ Sensitivity analysis or assessment of parameters effects on a model output when parameters are subject to variation ; b/ Uncertainty analysis, which enables taking into account errors in prominent parameters, and associating a global uncertainty confidence interval to the model output.

2- The comparison to the experiment, which consists in establishing two quality criteria to verify the model's accuracy. The first criterion consists in determining a validation threshold while comparing confidence intervals between predictions and measurements. The second criterion consists in verifying that the difference between measurement and calculation can be minimised using modification of the parameters in given proportions. This modification is obtained by applying a second order minimisation procedure.

3- The model improvement, which must take place when the model output is not valid according to the second quality criterion. This stage uses the results obtained during both minimisation procedure and sensitivity analysis.

This paper provides a description of the different tools used along with their application to two building models, one representing an experimental cell, another one representing a full scale building.

THE MODELS UNDER ANALYSIS

The first modelled building is the Etna test-cells, a real-size experimental building consisting of two semi-detachable cells set adjacent to each other. The modular configuration of one of the surrounding thermal guards (thermal guard n°3) makes it suitable for carrying out tests under natural or artificial climatic conditions (Fig.1).

This laboratory has a dual purpose: (1) to carry out thermal or aeraulic tests on components, while taking advantage of their semi-detached configuration to make real-time comparisons (natural climate) ; (2) to validate the models in CLIM2000 with respect to physical phenomena taken one by one.

During the experiment, the south wall of the cell n°2 was submitted to actual climate conditions and thermal guards were maintained at constant temperature. A purely convective heat source was switched on during the last 3 days of the 6-day experiment, according to a power step. The model involves 390 parameters, and the output analysed is the prediction of indoor air temperature corresponding to the cell n°2.

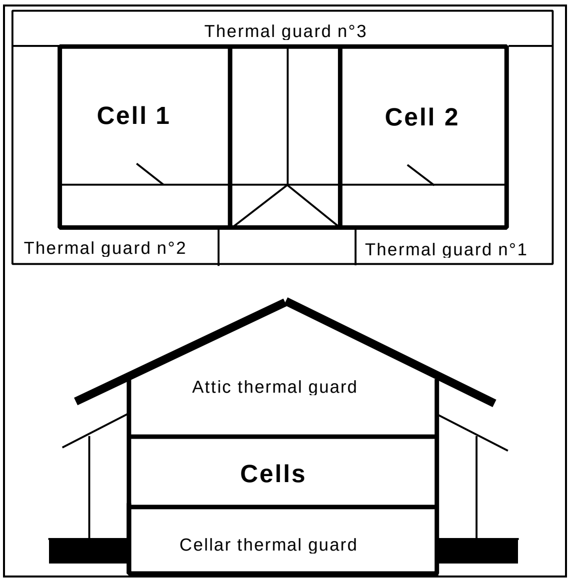

Fig.1/ The ETNA test cells

Experimental validation using test cells is invaluable for enhancing prediction power of models. However, these type of tests lack credibility among the wider audience. A complementary way for EDF is then to consider the CLIM2000 validation in a more global fashion.

The second modelled building is a house rented by EDF in order to treat experimental validation of full-scale buildings. This house - Valeriane - is a new cottage complying with 1989 French thermal regulations and is equipped with electric convectors. It is fully representative of all new dwelling units in France. This house is part of an allotment (urbanised area) rather free of nearby obstacles and with clear view especially to the south. It is a one-storey cottage with an attached garage.
During one week before the beginning of the experimental sequence, the thermostat of each heater was adjusted to obtain a 19°C constant temperature in each room. During the experimental sequence, from October 95 until May 96, the thermostat adjustments were never modified ; the general extraction unit was in operation ; there were no temperature control of heaters, no air mixing and no occupant inside ; the shutters were opened at 07 :00 a.m. and closed at 07 :00 p.m. each day.

The house is modelled with a simplified monozone model (i.e. one air volume representing the total volume of the house), and the model involves 540 parameters. The output analysed is the global simulated electrical power consumption during all of the experimental sequence.

DESCRIPTION OF TOOLS

1- Screening and Uncertainty analyses

Screening is a kind of sensitivity analysis in which all interactions and non-linearities are neglected. J.P.C Kleijnen [1995] defines screening as *'the search for the few really important factors among the great many N potentially important factors'*. Indeed, there are many parameters that potentially affect a typical process, but as a rule, only a small number of them are truly important.

In this study, two different pairs of screening/uncertainty analyses are presented : 1/ the exact differential analysis and its associated uncertainty analysis ; 2/ the group screening analysis and the Monte Carlo method. The first pair is deterministic, while the second is statistical.

Exact differential screening

This first kind of screening method is structured on the behaviour of the model for a base case scenario and involves differentiation of the equations describing the model with respect to a parameter. As the use of this rigorous method involves modifications of the code, this kind of screening analysis is not commonly applied to complex numerical models, because it can be very difficult to implement and it often requires very large amount of human and/or computer time.

Recently, a new capability of the CLIM2000 software program allows to calculate the first order derivatives of a dynamic result [Stangerup et al., 1997 ; Rahni et al., 1997]. The exact differential technique requires one single run of the model, which is very advantageous when considering large set of parameters.
In addition, derivatives allow to perform uncertainty analysis.
As we are considering first order approximation, the output uncertainty band can be assessed by :

$$\boxed{\Delta y(t) = \sum_i \Delta a_i \frac{\partial y(t)}{\partial a_i}}$$

where the quantity

$$\Delta a_i \frac{\partial y(t)}{\partial a_i}$$

gives the effect, on the output, of an error Δa_i on the time-invariant parameter a_i : for a dynamic model output, an effect is thus time-dependent.

Group screening and Monte Carlo method

Screening can also be performed by employing a statistical approach based on regression analysis and experimental designs. Although classical statistical screening methods are efficient for models involving a small number of parameters (around 50), they become unsuitable when this number reaches several hundreds. Then one must use group screening method, which allows to consider a large number of parameters.

This technique combines individual parameters into groups, and tests if these groups have a significant effect on the considered model output. Group screening is based on the following assumptions : 1/ parameters have independently, the same prior probability of being active ; 2/ an active parameter produces a non-zero change in the mean of the response ; 3/ we assume that the model can be approximated by main effects only ; 4/ a group effect is the sum of main effects corresponding to parameters which constitute the group.

Parameters included in non-significant groups are eliminated and new groups are constituted with the remaining parameters. The procedure continues until the number of remaining parameters is small enough for classical screening methods. At each step of the procedure, the model is replaced by the following first order polynomial function :

$$\boxed{y(t) = b_0(t) + \sum_{j=1}^{k} b_j(t) x_j + e(t)}$$

The *b* regression coefficients are fitted using ordinary least squares criterion and a set of simulation runs. The major advantage of this procedure is an easy implementation (for complex models), but it requires at each step $N \approx k+1$ model runs , becoming unsuitable when the number of parameters reaches several hundreds (for more details, see [Rahni et al., 1997]).

The uncertainty analysis is carried out by the *Monte Carlo* sampling strategy using the SPOP/PREP statistical pre- and post-processors of the LISA code [Saltelli et al., 1992]. This method gives a good assessment of the output uncertainty band (derived from the standard deviation), but requires a great number of simulations.

2- The minimisation procedure

The chosen procedure is the iterative Levenberg-Marquardt procedure [Eykhoff, 1979], which is a member of the Newton methods family. At each iteration of the descent algorithm, parameters are adjusted by means of the Newton direction *p*, as follow :

$$\boxed{\begin{aligned} a^{(k+1)} &= a^{(k)} + p^{(k)} \\ p^{(k)} &= -\left[He^{(k)} + \gamma^{(k)} I\right]^{-1} Gr^{(k)} \end{aligned}}$$

where γ is a small positive number, *He* and *Gr* being respectively the second derivative Hessian matrix and the gradient vector of the function to be minimised. In our case, this function is given by the quadratic criterion :

$$J(a) = \sum_{t=1}^{Mp} e^2(t,a)$$

where *e* is the difference $e(t,a) = \hat{y}(t,a) - y(t)$ between the assessed and the measured model output. In this particular case, an approximation of the gradient vector can be used by employing the first order matrix $M = \left[\frac{\partial e(t,a)}{\partial a}\right]$ only, as follows :

$$\boxed{Gr = 2Me}$$

The major handicap of Newton methods being the calculation of the Hessian matrix, a second approximation known as the Gauss-Newton technique can be made (only for quadratic criterion cases), which allows to obtain *He* more easily :

$$\boxed{He = 2MM^T}$$

Then, the capability of the CLIM2000 software program to perform exact derivative calculations makes if possible to conduct this optimisation procedure.

APPLICATION OF METHODOLOGY

1 - The model analysis

The goal at this stage is to select among a large set of parameters, the one to which models predictions are sensitive, and to predict the model output uncertainty that results from uncertainties on the prevailing parameters.

We define a parameter as a quantity that is time-invariant, whereas a model input has a dynamic behaviour. Thus thermo-physical and optical properties of building components, geometrical dimensions and site location constitute model parameters.

We first compared performances of both deterministic and statistical screening.

Screening

For the two models and by using the two screening methods, the same prominent parameters were exhibited. We selected 23 parameters (7%) from the 340 initial for the ETNA model, and 30 parameters (6%) from the initial 540 for the Valeriane model.

Using group screening, these results required 136 runs for the cell and 196 runs for the house, while differential screening was performed after one single run of the models.

Although the parameters effects were strongly dependent on time (Fig. 2 and 3), the exhibited parameters were all related to building components having the most important exchanged heat flow with the indoor air.

These components are :
1/ for the cell model : plasterboard and air gap of the north wall, concrete slab and insulation of the floor, and the south glazing;
2/ for the house model : insulation of the east and west walls, concrete and insulation of the floor, insulation of the ceiling, and the south glazing.

In addition, it appears that parameters were selected in relation with the heat storage capability of each component :
- for very resistive materials (like insulation), figure 2 indicates that the selected parameters are all member of the thermal resistance $R = e/ lS$, and have similar effects: conductivities (l) and surfaces *(S)* have same effects (module and direction), while thickness *(e)* acted in an opposite way;
- for materials with strong inertia (such as concrete), these parameters are those of the thermal capacity $C_{th} = eSrC_p$, and have similar effects also (Fig. 3): same direction and approximately same weight for thickness (*e*), surface (*S*), specific heat (C_p) and density (r) .

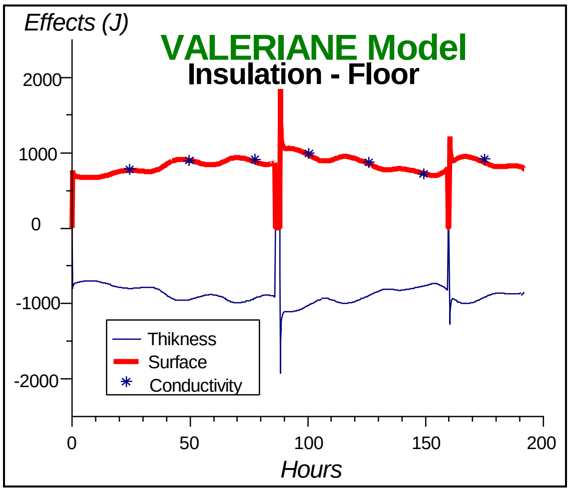

Fig.2/ Thermal resistance parameters effects

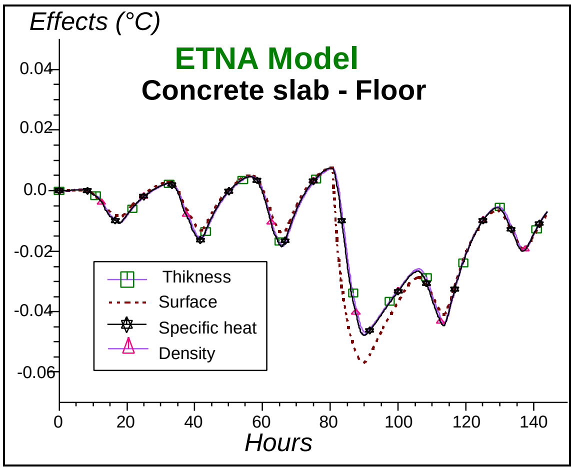

Fig.3/ Thermal capacity parameters effects

Uncertainty analysis

We applied the sampling strategy with 1000 runs on the prominent parameters. We supposed that each parameter was characterised by a uniform distribution law.

In addition, uncertainties of parameters were taken as follows:

$\pm$ 3% for thicknesses; $\pm$5% for surfaces; $\pm$10% for concrete slab thermo-physical parameters; $\pm$5% for the other thermo-physical parameters; $\pm$20% for convection heat transfer coefficients; $\pm$2% for tilt angles; $\pm$10% for azimuth angles.

Then the output uncertainty band assessed with this sampling method was compared with the one obtained from first order derivatives.

Results show a good agreement between the two approaches, probably due to the fact that the analysed models are weakly non-linear : consequently, a first order approximation makes it possible to compute output uncertainty band, with one single run.

The confidence intervals of both simulated consumption and simulated temperature are presented on figures 4 and 5.

2 - Comparison to experiment

The goal of this stage is to verify that the model can accurately predict the physical reality. Two quality criteria were then established :

First criterion

This criterion is based on the comparison of confidence intervals between predictions and measurements. When measurement error is not available or very low (as it is the case for the Etna and Valeriane models), the prediction uncertainty band must include 90% of the measurement at least, to say that the model is valid.

The confidence interval of the Valeriane model output is shown Figure 4 : the prediction overestimates the actual data with a rate of 5.4% and the width of the uncertainty band is around 20%.

In addition, the measured power is close to the lower limit : the model validity is not challenged by these results.

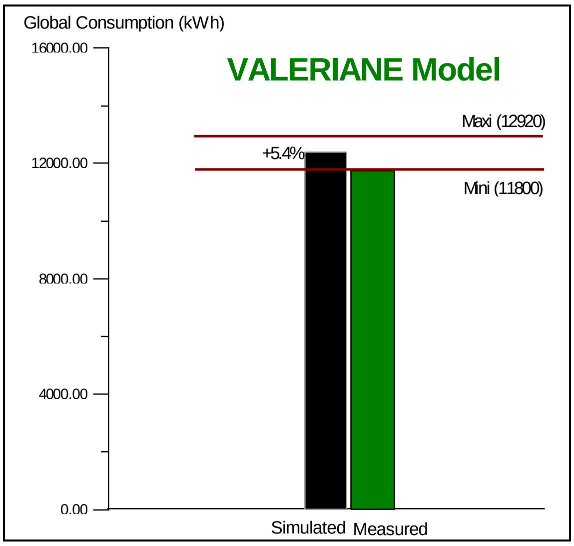

Fig.4/ Global power consumption - Valeriane model

However, for the Etna model, 63% of the points of actual data are outside the prediction uncertainty band : this ratio indicates that the model is non valid.

On figures 5 and 6, one can see that the air temperature is overestimated by the model when the power is switched on (at time = 81 hours).

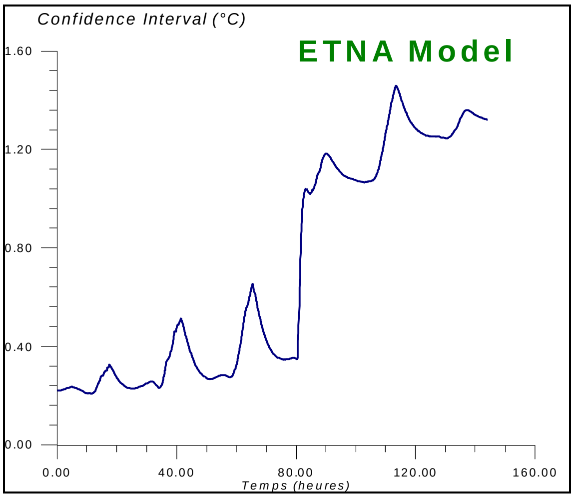

Fig.5/ Confidence Interval - Etna model

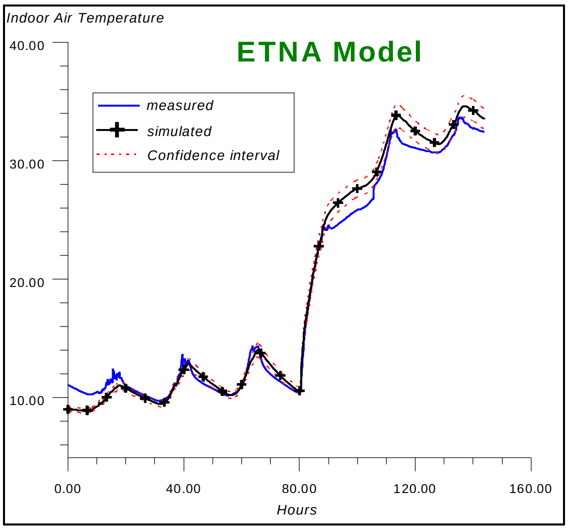

Fig.6/ Indoor air temperature - Etna model

In order to find out how to reduce this model discrepancy, the second quality criterion is applied.

Second criterion

This criterion consists in verifying that the model prediction can be improved using a modification of parameters as the result of a minimisation procedure : the Levenberg-Marquardt procedure.

As described in the screening section, influential parameters are all members either of a thermal resistance or a thermal capacitance and have similar effects. In a previous study, we have shown that parameters which constitute the same physical group (resistance or capacitance) are strongly correlated, these correlations making the identification job very hard or impossible to achieve.

For this reason, the authors consider that each group must be represented by a single parameter. This parameter is arbitrary chosen within the group.
On the basis of the 23 important parameters of the model, the procedure was then applied on the following parameters:

Tab.1/ The identified parameters

N°	Parameter
1	north wall : conductivity of air gap
2	floor : density of concrete slab
3	floor : internal heat transfer coefficient
4	floor :conductivity of insulation
5	south glazing : U-value

The authors have first considered an analytical test of the Levenberg-Marquardt procedure: parameters values are modified in an arbitrary way, and the authors verified that the procedure is able to identify reference values. This test was conclusive.

Then, the procedure was applied to the realistic configuration. Table 2 shows values given by the minimisation procedure. Expect for the convection heat transfer coefficient (parameter n°3), these values are in fact the result of either thermal resistances adjustment, or thermal capacities adjustment, explaining the large deviation between the reference case and the identified case.

The physical interpretation for these identified values is the following :
- the exchanged heat flows through the north wall are slightly reduced by increasing the global wall resistance (*via* the air gap thermal resistance);
- the exchanged heat flows through the south glazing are reduced by modifying the glazing U-value ;
- the exchanged heat through the floor towards outside the cell is increased by : 1/ increasing the convective flow and decreasing the storage effect for concrete; 2/ increasing the conductive flow for insulation.

Moreover, Table 3 indicates that the deviation between prediction and measure was thus significantly reduced.

Tab.2/ Values given by the minimisation procedure

N°	Base case	Identified values	Ratio (%)
1	0.287	0.16 ± 0.01	-44%
2	2060	1665 ± 40	-19%
3	9.09	19 ± 2	+110%
4	0.029	0.19 ± 0.01	+550%
5	3.47	2.5 ± 0.3	-29%

Tab.3/ Deviation Analysis

Deviation	measure – base case	Measure-new prediction
Mean	**-0.66**	-0.03
Standard deviation	**0.74**	0.31
Min	**-2.06**	-1.50
Max	**1.50**	1.18

On Figure 7, the three air temperatures - measured, at the base case, at the identified values- are plotted.

If experiment is accurately described by the model on the non-heating period and for the base case, a simulation defect is highlighted when the heat power is switched on, particularly at night-time (intervals 90h-103h and 115h-125h on Figures 7 and 8).

This fault seems to be explained by an inertia overestimation phenomenon. Indeed, the fact that heated and nocturnal periods occur after the heat storage within the concrete slab increases, entails a more important heat return. Then examination of Figures 7 and 8 confirms the identified values analysis.

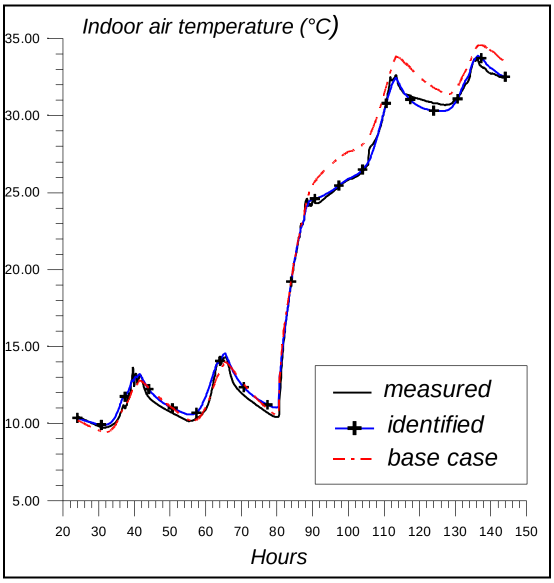

Fig.7/ Air temperatures : Measured - At the base case - At the identified values

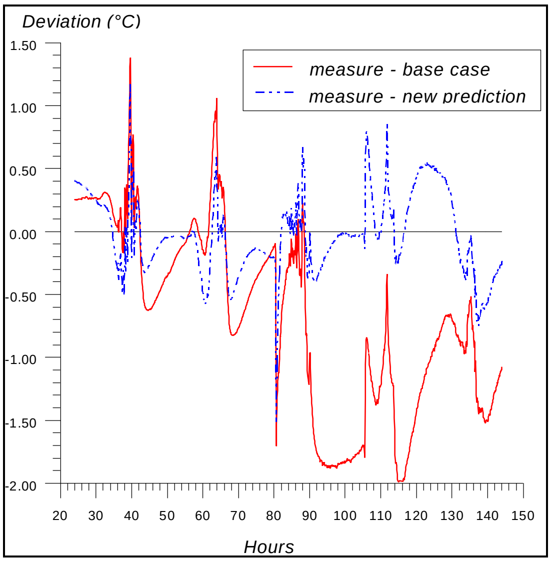

Fig.8/ Deviations : measure/base case and measure/new prediction

3 - The model improvement

The optimisation of the Etna model output shows that the model error is reduced when solar gains through the south glazing are diminished on the one hand, and when exchanged heat flows through the floor are increased on the other hand.

The fact that transmitted solar radiations are modelled as absorbed in totality by the floor slab, and that the identified value of the convection heat transfer coefficient is clearly non-realistic, indicates that one of the model improvement will be a new solar radiations distribution inside the cell (according to the wall surfaces for example)

CONCLUSIONS

This work has allowed to test different model improvement tools, by applying them to two building models.

At the close of this study, an important note concerning the capability of the CLIM2000 software program to perform exact derivative calculations came up : this advantage of the software make the sensitivity-uncertainty-optimisation work very accessible without increase of computer time.

We have shown that only a few parameters (around 7%) were really important for the two analysed model outputs. Error on these parameters will have non-negligible effects on the models predictions. Hence it is essential to have good estimates (reliable measure...) of these influential parameters in order to reduce outputs uncertainty bands.

The comparison between simulations and measurements has shown that 1/ for a global behaviour approach, as the case of global power consumption assessment, the building envelope model validity was not challenged; 2/ for a finer approach of phenomena, like dynamic temperature study, faults were exhibited. The minimisation procedure has allowed to see that these simulation errors were mainly related to an inertia overestimation phenomenon.

REFERENCES

P. Eykhoff, "System Identification", John Wiley & Sons, 1979

G. Guyon and N. Rahni, "Validation of a Building Thermal Model in CLIM2000 Simulation Software Using Full-Scale Experimental Data, Sensitivity Analysis and Uncertainty Analysis", CLIMA2000, 97

J. P. C. Kleijnen, "Sensitivity Analysis and Related Analysis, a Survey of Statistical Techniques", Research Memorandum, FEW 706, Faculty of Economics and Business Administration, Tilburg University, 1995

N. Rahni, N. Ramdani, Y. Candau and G. Guyon, "Exact Differential sensitivity analysis - Application to dynamic energy models developed on the CLIM 2000 software", EUROTHERM, Vol.10, pp.99-106, Mons (Belgium), October 8-10, 1997 (ISBN 92-827-5530-4)

N. Rahni, N. Ramdani, Y. Candau and P. Dalicieux, "Application Of Group Screening To Dynamic Buildings Energy Simulation Models", Journal of Statistical Computation & Simulation, Volume 57, Numbers 1-4, pp.14-16, Belgirate (Italy), September 25-27, 1995

A. Saltelli, T. Homma, "Statistical PRE Processor and Statistical POst Processor - Program description and User guides", Seminar on Sensitivity Analysis of Model Output, ISPN/DES/SESID Bureau of Fontenay-aux-Roses, Paris, 1992

P. Stangerup "Computation of Parameter Sensitivities in Non-Linear Dynamic Systems", EUROTHERM, Vol.10, pp.99-106, Mons (Belgium), October 8-10, 1997 (ISBN 2-84299-069-2)